Calculating Statistical Significance and Margins of Error Using American Community Survey Data

Montgomery County Census Workshop

Mark Goldstein

Maryland Department of Planning

March 15, 2011

OR …

MAKING IT EASY…TO USE ACS DATA!

ACS Sample Size for Maryland Housing Units

Year	Sample Size (final interviews)	Housing Unit Estimate	Percent
2005	**31,474**	**2,273,608**	**1.38%**
2006	**32,435**	**2,299,774**	**1.41%**
2007	**31,886**	**2,318,513**	**1.38%**
2008	**31,915**	**2,332,382**	**1.37%**
2009	**31,215**	**2,341,194**	**1.33%**
Total (2005-2009)			**6.87%**

Three-year sample size (2007-2009) = 4.1%
Five-year sample size = 6.9%
Decennial Census Long Form = 16.7%

Sampling Error & Standard Error

- **Sampling error** occurs when estimates are derived from a sample rather than a census (complete count) of the population.

- **Standard error** is an estimate of sampling error – how precise the survey estimates are to the true population you are trying to measure

Sampling Error & Margin of Error

- **Margin of Error** = standard error for a given confidence interval (typically 90 percent). A measure of the precision of the estimate at a given confidence interval

- Sampling error in the ACS is reported as the estimate “plus or minus” the margin of error

Margin of Error (MOE)

- **MOE = 1.645 * Standard error**
  where 1.645 is used for the 90 pct confidence interval
  (if want 95%, use 1.960; for 99% use 2.576)

- **Use the MOE to construct the Lower and Upper bounds around the estimate**

- **Lower Bound = (estimate – MOE)**
- **Upper Bound = (estimate + MOE)**

Median Household Income Estimate for Montgomery Co from the 2009 ACS

B19013. MEDIAN HOUSEHOLD INCOME IN THE PAST 12 MONTHS (IN 2009 INFLATION-ADJUSTED DOLLARS) - Universe: HOUSEHOLDS
Data Set: 2009 American Community Survey 1-Year Estimates
Survey: American Community Survey

NOTE. Although the American Community Survey (ACS) produces population, demographic and housing unit estimates, it is
Population Estimates Program that produces and disseminates the official estimates of the population for the nation, states,
estimates of housing units for states and counties.

For information on confidentiality protection, sampling error, nonsampling error, and definitions, see Survey Methodology.

	Montgomery County, Maryland	
	Estimate	Margin of Error
Median household income in the past 12 months (in 2009 inflation-adjusted dollars)	94,420	+/-2,347

Source: U.S. Census Bureau, 2009 American Community Survey

90 % C.I. = $94,420 +/- $2,347
= $92,073 to $96,767

90 Percent Confidence Interval

- There is a nine-out-of ten, or 90 % chance, that the interval contains the “true” value that you would have gotten from a full census

Why you should care about Margins of Error

- Lets you know how good the data is
- Saves you from drawing erroneous conclusions.
- Helps you decide how confident you can be about the assertions you make

First Example: Comparing Two Estimates

- If have two estimates, need to determine if the apparent differences are “real” (i.e. are statistically significant)

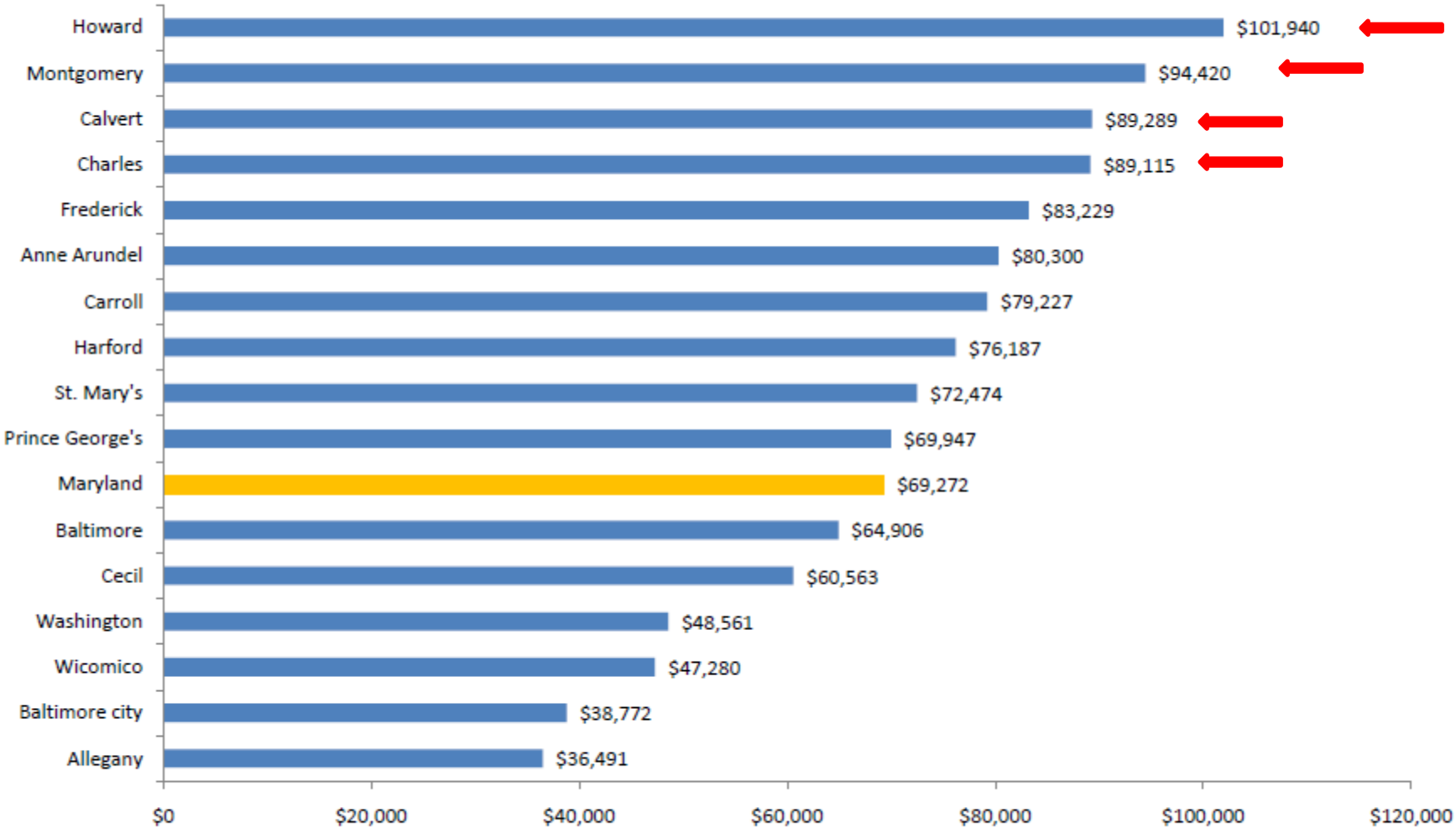
- Quick and dirty method is to “eye ball” whether the confidence intervals of the two estimates overlap

Comparing Two Estimates (the easy way)

- If the confidence intervals of two estimates do ***not*** overlap, then the difference between the two estimates are statistically significant

- If the confidence intervals of two estimates do overlap, then the difference between the two estimates may or may not be statistically significant (will need to test)

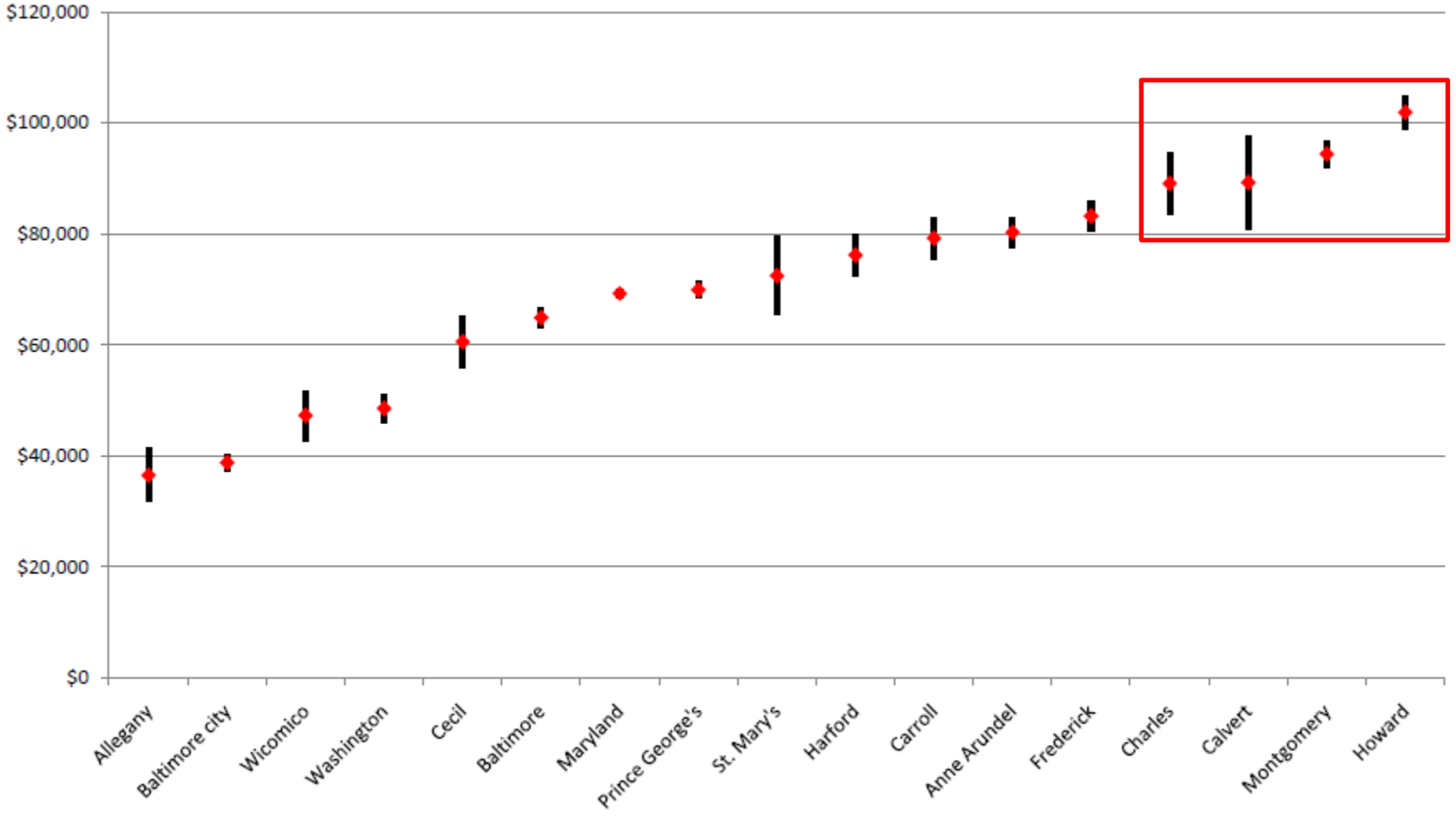
Median Household Income in Maryland and its Jurisdictions, 2009 (In 2009 Inflation-Adjusted Dollars) (Table GCT1901) *



* Apparent differences may not be statistically significant at the 90 percent confidence interval.

Prepared by the Maryland Department of Planning, from the 2009 ACS for the 16 counties covered (population of 65,000 or more), September 2010

Median Household Income in Maryland and its Jurisdictions, 2009 (In 2009 Inflation-Adjusted Dollars) (Table GCT1901) *



* Apparent differences may not be statistically significant at the 90 percent confidence interval.

Prepared by the Maryland Department of Planning, from the 2009 ACS for the 16 counties covered (population of 65,000 or more), September 2010

2009 Median Household Income Estimates

Jurisdiction	Median Income Estimate	Margin of error (MOE)	Lower Bound	Upper Bound
Howard	$101,940	$3,070	$98,870	$105,010
Montgomery	$94,420	$2,347	$92,073	$96,767
Calvert	$89,289	$8,465	$80,824	$97,754
Charles	$89,115	$5,488	$83,627	$94,603

Source: 2009 American Community Survey

Comparing Two Estimates

- Need to do a formal test of statistical significance if the confidence intervals do overlap

Testing Statistical Significance

1. Absolute value of Difference = ABS(X –Y)
2. SE(x) = MOEx/1.645
3. SE(y) = MOEy/1.645
4. SE(x-y) $\sqrt{[SE(X)]^2 + [SE(Y)]^2}$
5. MOE (x-y) = SE (x-y)*1.645
6. ABS(X-Y) <> MOE (x-y)

Testing Statistical Significance

1. If ABS(X-Y) > MOE(x-y), then the difference between the two estimates are statistically significant

2. If ABS(X-Y) < MOE (x-y), then the difference between the two estimates are NOT statistically significant

Go to Statistical Calculations Excel File! (significance test worksheet)

Thank you Dr. Lenny Gaines,

New York State Data Center

Testing a Difference Over Time

Montgomery County - Percent in Poverty

Age	ACS 2009		ACS 2008		ACS 2007	
	Est	Moe	Est	Moe	Est	Moe
LT 18	8.6%	1.7%	7.7%	1.6%	4.7%	1.4%
65+	6.3%	1.3%	5.7%	1.0%	5.3%	1.4%

Go to Statistical Calculations Excel File! (significance test worksheet)

Second Example – Calculate MOE of the sum of two or more estimates

Top Six Places for Foreign Born in Montgomery County (2005-2009)

Place	Estimate	Margin of Error
Silver Spring	26,404	1,591
Wheaton-Glenmont	25,217	1,551
Gaithersburg	21,505	1,321
Rockville	20,306	1,031
Germantown	20,042	1,597
Aspen Hill	18,747	1,486
TOTAL	**132,221**	

Second Example – Margin of Error of a Sum

1. **Calculate the standard error of the sum:**

SE $(x_1+x_2+x_3)$=

$$= \sqrt{[MOE(x1)/1.645]^2 + [MOE(x2)/1.645]^2 + [MOE_{x3}/1.645]^2}$$

2. **Calculate the Margin of error of the sum:**

$$MOE(x_1+x_2+x_3) = SE(x_1+x_2+x_3)*1.645$$

Go to Statistical Calculations Excel File! (StdErrSum3ormore worksheet)

Second Example – Calculate MOE of the sum of two or more estimates

Top Six Places for Foreign Born in Montgomery County (2005-2009)

Place	Estimate	Margin of Error
Silver Spring	26,404	1,591
Wheaton-Glenmont	25,217	1,551
Gaithersburg	21,505	1,321
Rockville	20,306	1,031
Germantown	20,042	1,597
Aspen Hill	18,747	1,486
TOTAL	**132,221**	**3,536**

Third Example – Margin of Error of a Proportion

(e.g., What proportion of Montgomery's foreign born population are in the top six places?)

X = number of foreign born in top six places
Y = number of foreign born in Montgomery County

i.e. X is a subset of Y

$$SE(x/y) = 1/Y * \sqrt{[MOEx/1.645]^2 - [x^2/y^2]*[MOEy/1.645]^2}$$

OR

$$SE(x/y) = 1/Y * \sqrt{[MOEx/1.645]^2 + [x^2/y^2]*[MOEy/1.645]^2}$$

$$MOE\ (x/y) = SE(x/y)*1.645$$

Go to Statistical Calculations Excel File! (StdErr of a proportion worksheet)

Fourth Example – Margin of Error of a Ratio

(e.g. What is the ratio of median income of foreign born to native born households in Montgomery Co.

X = median income of foreign born households
Y = median income of native born households

i.e., X is NOT a subset of Y

$$SE(x/y) = 1/Y * \sqrt{[MOEx/1.645]^2 + [x^2/y^2] * [MOEy/1.645]^2}$$

MOE (x/y) = SE(x/y)*1.645

Go to Statistical Calculations Excel File! (Std Err of a ratio worksheet)

Thank You…..

Hope we made it (somewhat) easier

STATE DATA CENTER POPULAR LINKS

Zip Code Maps
MD Statistical Handbook
Population Estimates
Median Household Income Estimates
Projection Profiles
Parcel PFA Data

STATE DATA CENTER QUICK LINKS

Census 2010 Activities
Projections
Estimates
Building Permit Data
American Community Survey
Job & Personal Income Data
Map & GIS Products
Economic & Agricultural Census
2010 Statistical Abstract

American Community Survey

The American Community Survey (ACS) is an ongoing, nationwide, monthly survey covering approximately three-million households each year and is now the source for all of the socioeconomic data that was formerly obtained from the decennial census. ACS data is available for three types of releases depending on population size: annually, three-year periods and five-year periods.

Annual ACS data are available for all geographies (e.g., U.S., states, counties, municipalities, places) with a population of 65,000 or more.

Three-year period ACS estimates are available for all geographies with a population of 20,000 or more.

Five-year period ACS estimates are available for all geographies down to census tracts and block groups.

Data and Reports

- Cautions, Documentation and Guidelines on Using ACS Data
- ACS Annual Data : 2009 *new* || 2008 || 2007 || 2006 || 2005 || 2004 || 2003 || 2002
- ACS Three-year Data : 2007 to 2009 *new* || 2006 to 2008 || 2005 to 2007
  Socioeconomic Characteristics for Maryland's Jurisdictions and Places : 2007 to 2009 *new* || 2006 to 2008 || 2005 to 2007
  Census Transportation Planning Products (CTPP) : 2006 to 2008
- ACS Five-year Data : 2005 to 2009 *new*

Calculations of Statistical Significance & MOEs of Combinations of ACS Data

- To Calculate If Two Estimates are Statistically Significant
- Spreadsheet for Calculating Margins of Error and Statistical Significance for Sums, Proportions and Ratios